W382–W388 *Nucleic Acids Research, 2014, Vol. 42, Web Server issue* *Published online 15 May 2014*
doi: 10.1093/nar/gku438

COGNAC: a web server for searching and annotating hydrogen-bonded base interactions in RNA three-dimensional structures

Mohd Firdaus-Raih[1,2,*], Hazrina Yusof Hamdani[1], Nurul Nadzirin[1], Effirul Ikhwan Ramlan[3], Peter Willett[4] and Peter J. Artymiuk[5]

[1]School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Malaysia, [2]Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Malaysia, [3]Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia, [4]Information School, University of Sheffield, Western Bank, Sheffield S10 2TN, UK and [5]Department of Molecular Biology and Biotechnology, Krebs Institute, University of Sheffield, Western Bank, Sheffield S10 2TN, UK

Received March 17, 2014; Revised April 18, 2014; Accepted May 3, 2014

ABSTRACT

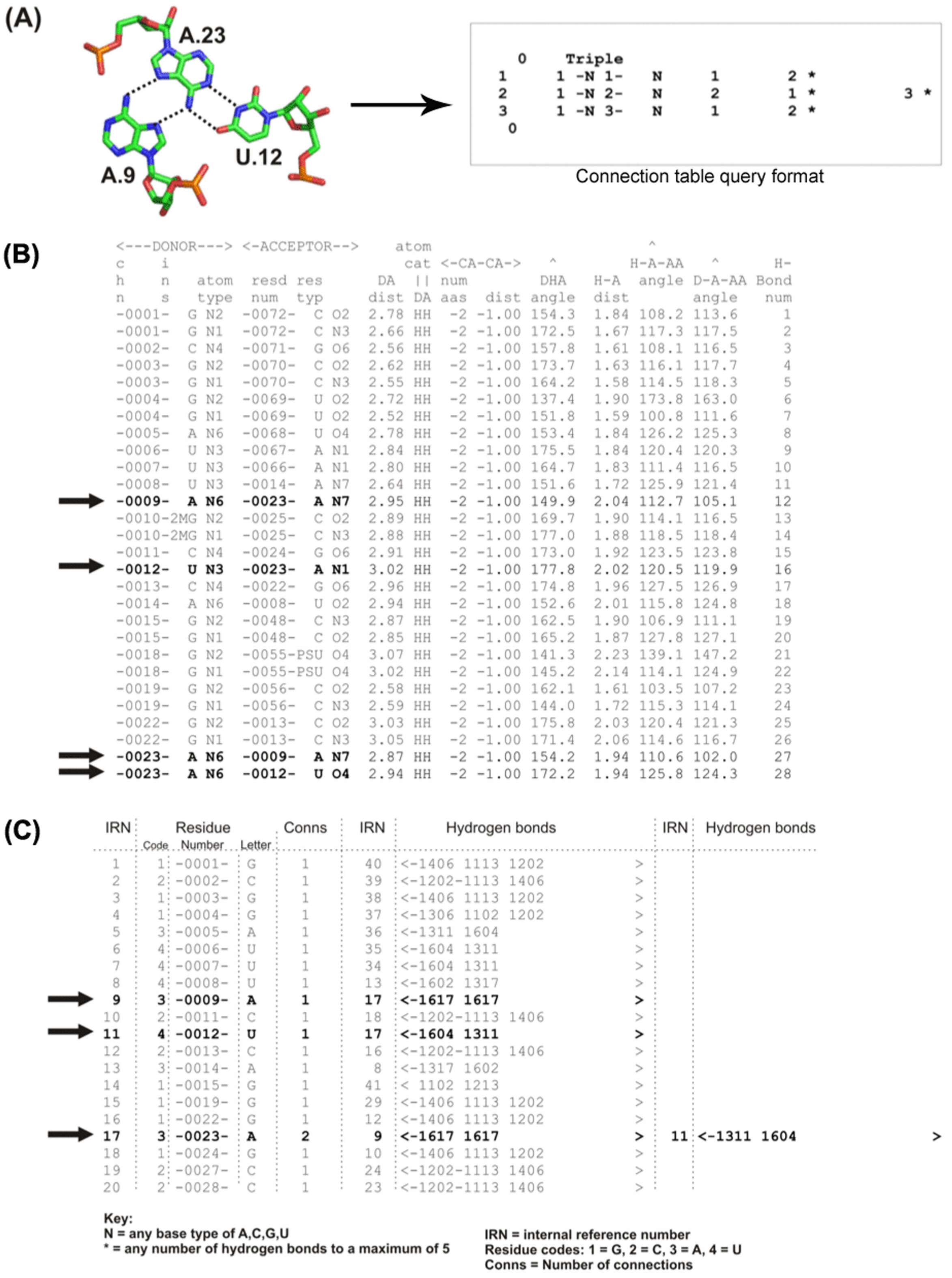
Hydrogen bonds are crucial factors that stabilize a complex ribonucleic acid (RNA) molecule's three-dimensional (3D) structure. Minute conformational changes can result in variations in the hydrogen bond interactions in a particular structure. Furthermore, networks of hydrogen bonds, especially those found in tight clusters, may be important elements in structure stabilization or function and can therefore be regarded as potential tertiary motifs. In this paper, we describe a graph theoretical algorithm implemented as a web server that is able to search for unbroken networks of hydrogen-bonded base interactions and thus provide an accounting of such interactions in RNA 3D structures. This server, COGNAC (COnnection tables Graphs for Nucleic ACids), is also able to compare the hydrogen bond networks between two structures and from such annotations enable the mapping of atomic level differences that may have resulted from conformational changes due to mutations or binding events. The COGNAC server can be accessed at http://mfrlab.org/grafss/cognac.

INTRODUCTION

The three-dimensional (3D) structures of complex ribonucleic acid (RNA) molecules are as crucial for their function as they are for proteins. The contribution by networks of hydrogen-bonded interactions towards structural stabilization have been well studied in RNA structures (1–5). Efforts have been made to annotate the 3D structures of RNA in order to analyze their functions. The majority of such structural annotations have focused on the alignment of RNA 3D structures to detect similarities in folding and sub-folding (6–8) and these approaches are not unlike those used in protein fold comparisons where an often used method involves superpositions of the atoms forming the structures' backbones. Other programs such as NASSAM (9) and WebFR3D (10) analyze the spatial arrangements of the RNA bases in order to identify motifs. Lescoute and Westhof had previously explored the annotation of interaction networks for three-way junctions in folded RNAs (11). This annotation effort involved the mapping of base–base interaction networks using the nomenclature system previously proposed by Leontis and Westhof (12) that further complemented the earlier work by Gutell *et al.* (13). In the literature, investigations that specifically focused on analyzing the interactions of base triples (example in Figure 1A, left panel) had been carried out by Abu Almakarem *et al.* (14) using FR3D (15) and Firdaus-Raih *et al.* (16,17) using the computer program NASSAM (9). However, these previous studies have not specifically explored regions of the RNA 3D structures that involve tight clusters of interacting bases or the occurrences of unbroken networks of hydrogen-bond-mediated base interactions.

The hydrogen-bonded base interaction networks within an RNA structure can be thought of as two-dimensional (2D) networks. Previously, Gan *et al.* (18) described the use of planar graphs to represent RNA structural topology. In a similar way, the hydrogen-bonded interactions between bases can also be represented via the use of tree graphs. In such a tree representation, hydrogen bonds can be represented by the graph's edges and these bonds can be either single, bifurcated or multiple interactions. Labels can be added to these graphs, where each node can represent one of the four bases or simply a 'wild-card' to represent any of

[*]To whom correspondence should be addressed. Tel: + 603 8921 5961; Fax: + 603 8925 2698; Email: firdaus@mfrlab.org



```
   0     Triple
1     1 -N 1-    N     1        2 *
2     1 -N 2-    N     2        1 *           3 *
3     1 -N 3-    N     1        2 *
 0
```

Connection table query format

```
<---DONOR---> <-ACCEPTOR-->    atom                              ^
c    i                          cat <-CA-CA->   ^          H-A-AA   ^     H-
h    n   atom  resd res       DA  || num          DHA  H-A  angle D-A-AA Bond
n    s   type  num  typ      dist DA aas   dist angle  dist        angle   num
-0001-  G N2  -0072-  C O2  2.78 HH  -2 -1.00 154.3  1.84 108.2 113.6     1
-0001-  G N1  -0072-  C N3  2.66 HH  -2 -1.00 172.5  1.67 117.3 117.5     2
-0002-  C N4  -0071-  G O6  2.56 HH  -2 -1.00 157.8  1.61 108.1 116.5     3
-0003-  G N2  -0070-  C O2  2.62 HH  -2 -1.00 173.7  1.63 116.1 117.7     4
-0003-  G N1  -0070-  C N3  2.55 HH  -2 -1.00 164.2  1.58 114.5 118.3     5
-0004-  G N2  -0069-  U O2  2.72 HH  -2 -1.00 137.4  1.90 173.8 163.0     6
-0004-  G N1  -0069-  U O2  2.52 HH  -2 -1.00 151.8  1.59 100.8 111.6     7
-0005-  A N6  -0068-  U O4  2.78 HH  -2 -1.00 153.4  1.84 126.2 125.3     8
-0006-  U N3  -0067-  A N1  2.84 HH  -2 -1.00 175.5  1.84 120.4 120.3     9
-0007-  U N3  -0066-  A N1  2.80 HH  -2 -1.00 164.7  1.83 111.4 116.5    10
-0008-  U N3  -0014-  A N7  2.64 HH  -2 -1.00 151.6  1.72 125.9 121.4    11
-0009-  A N6  -0023-  A N7  2.95 HH  -2 -1.00 149.9  2.04 112.7 105.1    12
-0010-2MG N2  -0025-  C O2  2.89 HH  -2 -1.00 169.7  1.90 114.1 116.5    13
-0010-2MG N1  -0025-  C N3  2.88 HH  -2 -1.00 177.0  1.88 118.5 118.4    14
-0011-  C N4  -0024-  G O6  2.91 HH  -2 -1.00 173.0  1.92 123.5 123.8    15
-0012-  U N3  -0023-  A N1  3.02 HH  -2 -1.00 177.8  2.02 120.5 119.9    16
-0013-  C N4  -0022-  G O6  2.96 HH  -2 -1.00 174.8  1.96 127.5 126.9    17
-0014-  A N6  -0008-  U O2  2.94 HH  -2 -1.00 152.6  2.01 115.8 124.8    18
-0015-  G N2  -0048-  C N3  2.87 HH  -2 -1.00 162.5  1.90 106.9 111.1    19
-0015-  G N1  -0048-  C O2  2.85 HH  -2 -1.00 165.2  1.87 127.8 127.1    20
-0018-  G N2  -0055-PSU O4  3.07 HH  -2 -1.00 141.3  2.23 139.1 147.2    21
-0018-  G N1  -0055-PSU O4  3.02 HH  -2 -1.00 145.2  2.14 114.1 124.9    22
-0019-  G N2  -0056-  C O2  2.58 HH  -2 -1.00 162.1  1.61 103.5 107.2    23
-0019-  G N1  -0056-  C N3  2.59 HH  -2 -1.00 144.0  1.72 115.3 114.1    24
-0022-  G N2  -0013-  C O2  3.03 HH  -2 -1.00 175.8  2.03 120.4 121.3    25
-0022-  G N1  -0013-  C N3  3.05 HH  -2 -1.00 171.4  2.06 114.6 116.7    26
-0023-  A N6  -0009-  A N7  2.87 HH  -2 -1.00 154.2  1.94 110.6 102.0    27
-0023-  A N6  -0012-  U O4  2.94 HH  -2 -1.00 172.2  1.94 125.8 124.3    28
```

```
IRN      Residue          Conns   IRN      Hydrogen bonds          IRN  Hydrogen bonds
    Code  Number  Letter
  1   1  -0001-   G       1      40  <-1406 1113 1202         >
  2   2  -0002-   C       1      39  <-1202-1113 1406         >
  3   1  -0003-   G       1      38  <-1406 1113 1202         >
  4   1  -0004-   G       1      37  <-1306 1102 1202         >
  5   3  -0005-   A       1      36  <-1311 1604              >
  6   4  -0006-   U       1      35  <-1604 1311              >
  7   4  -0007-   U       1      34  <-1604 1311              >
  8   4  -0008-   U       1      13  <-1602 1317              >
  9   3  -0009-   A       1      17  <-1617 1617              >
 10   2  -0011-   C       1      18  <-1202-1113 1406         >
 11   4  -0012-   U       1      17  <-1604 1311              >
 12   2  -0013-   C       1      16  <-1202-1113 1406         >
 13   3  -0014-   A       1       8  <-1317 1602              >
 14   1  -0015-   G       1      41  < 1102 1213              >
 15   1  -0019-   G       1      29  <-1406 1113 1202         >
 16   1  -0022-   G       1      12  <-1406 1113 1202         >
 17   3  -0023-   A       2       9  <-1617 1617              >   11 <-1311 1604            >
 18   1  -0024-   G       1      10  <-1406 1113 1202         >
 19   2  -0027-   C       1      24  <-1202-1113 1406         >
 20   2  -0028-   C       1      23  <-1202-1113 1406         >
```

Key:
N = any base type of A,C,G,U
*** = any number of hydrogen bonds to a maximum of 5**

IRN = internal reference number
Residue codes: 1 = G, 2 = C, 3 = A, 4 = U
Conns = Number of connections

Figure 1. (**A**) Base triple interaction example A9.A23.U12 from the yeast tRNA^{Phe} structure (PDB ID: 6tna, left panel) and it's corresponding representation in a connection table (right panel). (**B**) Partial sample of HBPRED output file with the four hydrogen bonds for the triple in (A) highlighted in bold. (**C**) Partial sample of the connection table file after conversion from the format in (B) with the interactions for the AAU triple highlighted in bold. The highlighted data in (C) can be read as 'base A9 (IRN = 9) is connected to one other base, A23 (IRN = 17), by two hydrogen bonds; base U12 (IRN = 11) is connected to one other base, A23, by two H-bonds; base A23 is connected to two other bases A9 and U12 by two hydrogen bonds each'. '−1604 1311' can be read as 'A23 N6 (1 = nitrogen) donor to U12 O4 (0 = oxygen) acceptor and U12 N3 donor to A23 N1 acceptor'.

the bases. The edges can also be similarly labeled in such a way that an edge can represent a specific hydrogen bond, a particular number of hydrogen bonds or any number of hydrogen bonds. The information within these labeled graphs can be contained within the tabular data structure of a connection table (Figure 1A, right panel). A graph theoretical algorithm can then be employed to study relationships between the connection tables (19).

As various structure-dependent functional RNA molecules are discovered and more RNA structures become available in the Protein Data Bank (PDB) (20), a computational tool that can search, annotate and provide comparisons between large numbers of structures is expected to be a useful utility for RNA biologists in general. In this article, we describe a server that is able to map hydrogen bonds between RNA bases and from this accounting of such interactions, allow for the annotation of RNA base interaction clusters that may potentially be structural motifs; in addition, it may also yield potential insights regarding the functional or structure stabilization roles for clusters of interacting bases.

PROGRAMS AND METHODS

The COGNAC (COnnection tables Graphs for Nucleic ACids) server consists of three major components: the hydrogen bond data generating program, the graph theoretical COGNAC algorithm and the visual analysis component. The hydrogen bond generating program, HBPRED (Hydrogen Bonding Predictor) , calculates the formation of a hydrogen bond between donor and acceptor atoms using the parameters previously used in HBPLUS (21). The HBPRED outputs (Figure 1B) are then converted into a connection table format (Figure 1C) that is searchable by the main COGNAC program. HBPRED generates the hydrogen bonding data for user provided structures as well as for structures in the internal database that were sourced from the PDB. An example of a base triple is provided in Figure 1A where a base triple with four hydrogen bonds is part of a transfer RNA structure (PDB ID: 6tna). The HBPRED program generates a list of all possible hydrogen bond interactions in the structure (Figure 1B) which is then converted into a connection table (Figure 1D). A query in connection table format for the triple (Figure 1B) that is processed through COGNAC is then able to identify the subgraph of the query (Figure 1A, right panel) in the larger connection table for the whole structure (Figure 1C).

The COGNAC program is a graph theoretical implementation of the Ullmann subgraph isomorphism (22) that solves the problem of matching a query connection table graph representation to its matching subgraph. The query submission interface allows for the searching and annotation of structures within an internal database or for user provided RNA structures (Figure 2A). The base interaction possibilities for COGNAC search patterns were enumerated and are represented as tree graphs that range from single possibilities for pairs and triples to six possible unique interaction arrangements for sextuples (Figure 2B). The query for a COGNAC search consists of the bases represented as the graph nodes and the edges representing the number of connections (Figure 1B), which in this case refers to the number of hydrogen bonding interactions to each base. The number of hydrogen bonds represented by the edges can be left unspecified with a minimum of one bond. The nodes representing the bases can be pre-determined as a specific base or left as a wildcard base representing any of the base options specified in the base library, which includes several modified bases such as pseudouracil.

INPUT AND OUTPUT

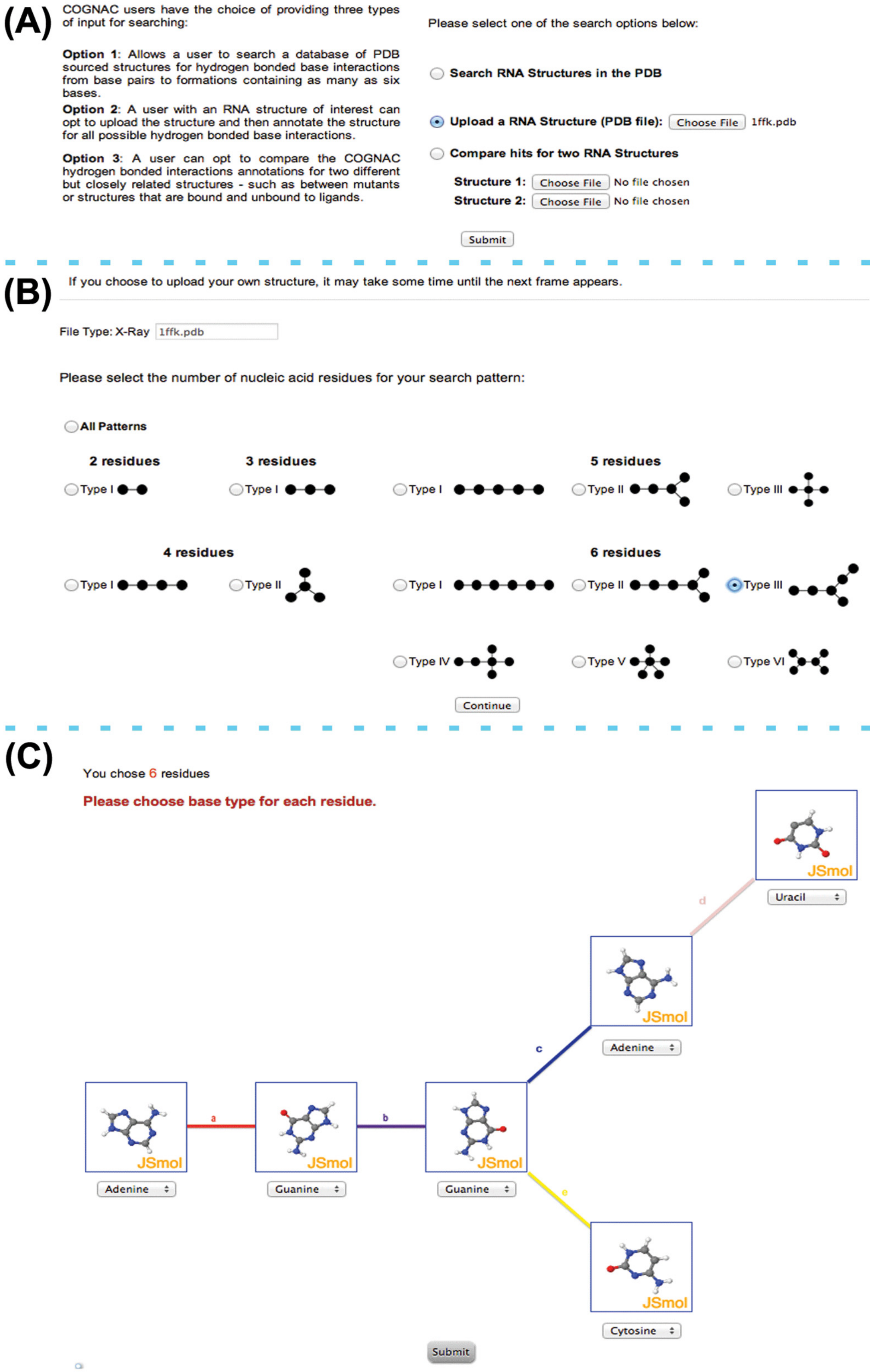
Input description

The COGNAC query interface allows a user to submit three types of search options (Figure 2A). The first option enables a user to search a pre-set database of RNA chain containing structures that were sourced from the PDB. The database is updated on a monthly basis with the current version reported in this paper compiled on 12 March 2014 and contains 1778 files with structure resolutions >3.5 Å. The second option allows a user to upload a PDB formatted structure of interest while the third search option provides the user with the capability to search two PDB formatted structures simultaneously, thus allowing for the differences between the hydrogen bonds formed within the two structures to be investigated. For structures that were generated using nuclear magnetic resonance spectroscopy (NMR) data, the user will be prompted to select only one model for analysis at a time. The program will also automatically remove all hydrogen atoms in NMR structures and thus these atoms will no longer be visible when the hits are visualized using either the Jmol or JSmol viewers.

Once the initial type of search has been selected, a new interface appears which provides the user with the option of either (i) searching all patterns of base pairs to base sextuples that have been enumerated or (ii) searching for a specific base-to-base connection that ranges from a pair to a sextuple (Figure 2B). Upon selection of the type of representative graph(s) for the connection table, another interface appears prompting the user to provide a specific base as the graph's node or to leave the base as a wildcard option (Figure 2C). This interface uses JSmol (http://wiki.jmol.org) windows to enable users to orientate the bases in order to facilitate a visual aided approach toward envisioning possible hydrogen bonding interactions that may occur. However, the orientations of the bases in the JSmol windows are not used as a parameter for the search because only information on the base type is required for the searches.

Output description

COGNAC searches for all three input options will first output a listing of the number of matches to a particular type of interaction (tree) queried. The raw output is filtered for redundant hits that arise due to the searches being independent of sequence order thus resulting in the same hits being retrieved for the same interaction for the different sequence order possibilities. The filtered output is presented to the user as a summary of the number of hits for a particular tree query or queries. Selecting the output for a specific tree graph then results in a list that provides information regarding the structure in which a match is found and a listing of the hydrogen bonds involved in a particular match.



(A)

COGNAC users have the choice of providing three types of input for searching:

Option 1: Allows a user to search a database of PDB sourced structures for hydrogen bonded base interactions from base pairs to formations containing as many as six bases.

Option 2: A user with an RNA structure of interest can opt to upload the structure and then annotate the structure for all possible hydrogen bonded base interactions.

Option 3: A user can opt to compare the COGNAC hydrogen bonded interactions annotations for two different but closely related structures - such as between mutants or structures that are bound and unbound to ligands.

Please select one of the search options below:

○ **Search RNA Structures in the PDB**

● **Upload a RNA Structure (PDB file):** Choose File 1ffk.pdb

○ **Compare hits for two RNA Structures**

Structure 1: Choose File No file chosen

Structure 2: Choose File No file chosen

(B)

If you choose to upload your own structure, it may take some time until the next frame appears.

File Type: X-Ray 1ffk.pdb

Please select the number of nucleic acid residues for your search pattern:

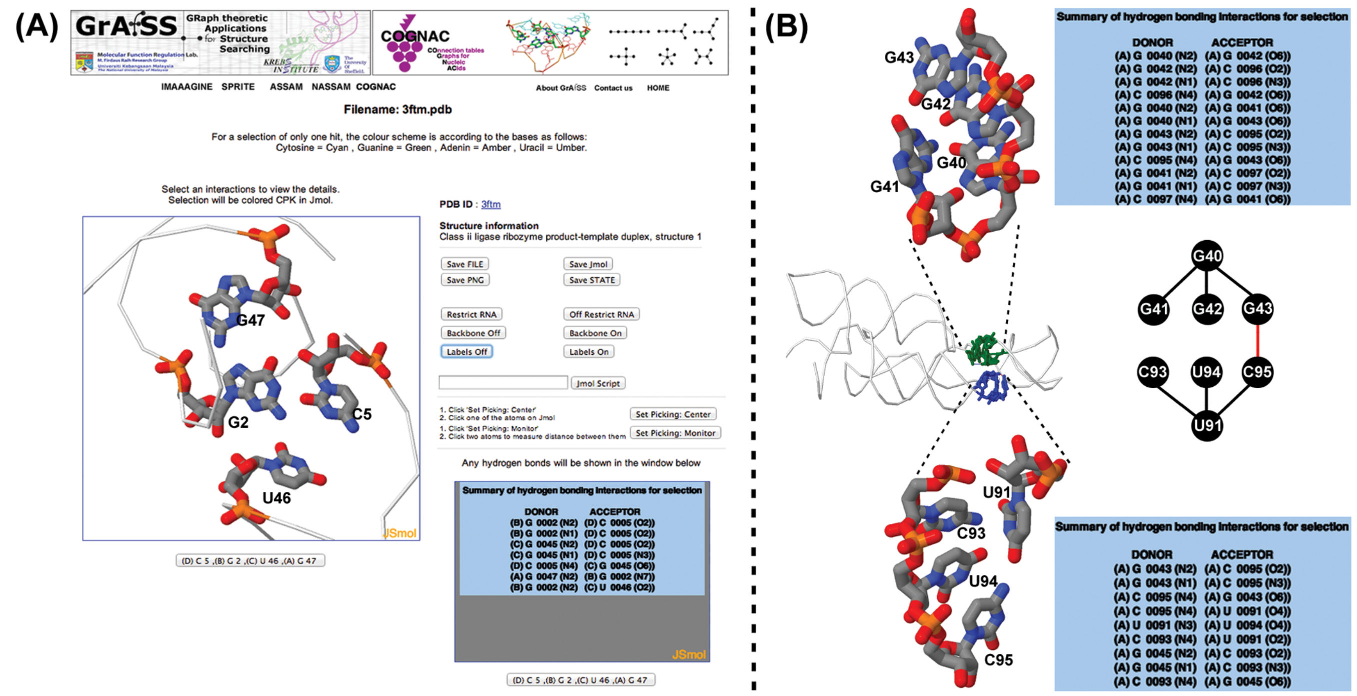
○ **All Patterns**

(C)

You chose 6 residues

Please choose base type for each residue.

Figure 2. Input interface for COGNAC where: (**A**) users can initially select between three types of search options; the example presented shows the option to upload a user provider query being selected. (**B**) A follow-up interface proceeding the initial options prompts users to select the type of interaction to search for based on the representative tree graphs for the interaction of interest; the example presented shows the selection of a sextuple arrangement. (**C**) An additional search specification option where users can define the bases for the query involved in the sextuple arrangement that has been selected in (**B**); users have the option of having the query design facilitated by a mini JSmol window depicting the molecular structure for each base.


(A)
GrAfSS GRaph theoretic Applications for Structure Searching
Molecular Function Regulation Lab.
KREBS INSTITUTE
COGNAC COnnection tables Graphs for Nucleic ACids
IMAAAGINE SPRITE ASSAM NASSAM COGNAC
About GrAfSS Contact us HOME
G47
G2
C5
U46
JSmol
Save FILE
Save Jmol
Save PNG
Save STATE
Restrict RNA
Off Restrict RNA
Backbone Off
Backbone On
Labels Off
Labels On
Jmol Script
Set Picking: Center
Set Picking: Monitor
(D) C 5 ,(B) G 2 ,(C) U 46 ,(A) G 47
(B)
G43
G42
G40
G41
G40
G41
G42
G43
C93
U94
C95
U91
U91
C93
U94
C95

Filename: 3ftm.pdb

For a selection of only one hit, the colour scheme is according to the bases as follows:
Cytosine = Cyan , Guanine = Green , Adenin = Amber , Uracil = Umber.

Select an interactions to view the details.
Selection will be colored CPK in Jmol.

PDB ID : 3ftm

Structure information
Class ii ligase ribozyme product-template duplex, structure 1

1. Click 'Set Picking: Center'
2. Click one of the atoms on Jmol

1. Click 'Set Picking: Monitor'
2. Click two atoms to measure distance between them

Any hydrogen bonds will be shown in the window below

Summary of hydrogen bonding interactions for selection

DONOR	ACCEPTOR
(B) G 0002 (N2)	(D) C 0005 (O2))
(B) G 0002 (N1)	(D) C 0005 (O2))
(C) G 0045 (N2)	(D) C 0005 (O2))
(C) G 0045 (N1)	(D) C 0005 (N3))
(D) C 0005 (N4)	(C) G 0045 (O6))
(A) G 0047 (N2)	(B) G 0002 (N7))
(B) G 0002 (N2)	(C) U 0046 (O2))

Summary of hydrogen bonding interactions for selection

DONOR	ACCEPTOR
(A) G 0040 (N2)	(A) G 0042 (O6))
(A) G 0042 (N2)	(A) C 0096 (O2))
(A) G 0042 (N1)	(A) C 0096 (N3))
(A) C 0096 (N4)	(A) G 0042 (O6))
(A) G 0040 (N2)	(A) G 0041 (O6))
(A) G 0040 (N1)	(A) G 0043 (O6))
(A) G 0043 (N2)	(A) C 0095 (O2))
(A) G 0043 (N1)	(A) C 0095 (N3))
(A) C 0095 (N4)	(A) G 0043 (O6))
(A) G 0041 (N2)	(A) C 0097 (O2))
(A) G 0041 (N1)	(A) C 0097 (N3))
(A) C 0097 (N4)	(A) G 0041 (O6))

Summary of hydrogen bonding interactions for selection

DONOR	ACCEPTOR
(A) G 0043 (N2)	(A) C 0095 (O2))
(A) G 0043 (N1)	(A) C 0095 (N3))
(A) C 0095 (N4)	(A) G 0043 (O6))
(A) C 0095 (N4)	(A) U 0091 (O4))
(A) U 0091 (N3)	(A) U 0094 (O4))
(A) C 0093 (N4)	(A) U 0091 (O2))
(A) G 0045 (N2)	(A) C 0093 (O2))
(A) G 0045 (N1)	(A) C 0093 (N3))
(A) C 0093 (N4)	(A) G 0045 (O6))

Figure 3. Examples of COGNAC output: (**A**) the molecular visualization interface using JSmol (or Jmol) with basic information on the structure, external links to the PDB entry and a listing of the hydrogen bonds involved in the interaction network which in this case is a base quadruple; (**B**) two matches for a type II quadruple (refer to Figure 2B) search tree query in a lysine riboswitch structure—a GGGG quadruple (center panel, in green) and a CUCU quadruple (center panel, in blue) that upon further scrutiny is interconnected (denoted by red line) to form an octuple base interaction (center panel).

A user can then select a particular hit for further inspection using the built-in molecular visualization interfaces (Figure 3). Molecular visualization can be carried out using either the Jmol or JSmol (http://wiki.jmol.org) options. Jmol, a Java plugin, is noticeably more responsive at handling large structures such as ribosomal RNA when compared to JSmol. However, some operating systems have been known to block the use of Java due to security concerns and Jmol use may also be inconvenient to some users due to the requirement of needing Java installed for the browser. The JSmol viewers on the other hand will work directly inside the browsers without any additional installation requirements. Both the Jmol and JSmol windows allow for in browser analysis of the interactions and a further interaction by interaction inspection of the hits. Each molecular visualization window is accompanied by a detailed listing of the hydrogen bond interactions computed by HBPRED (Figure 3). Additionally, the hydrogen bonds are by default drawn into the visual display for each hit with the respective bond lengths labeled. The server also generates the 2D structure representations from the HBPRED data using the RNAfold program from the ViennaRNA Package (23). For the comparison of similar structures, a listing that highlights the differences between two structures is provided with a further visualization option of comparing the structures in adjacent synchronized Jmol or JSmol windows.

DISCUSSION AND CASE STUDIES

Currently, several web servers such as WebFR3D (10), NASSAM (9), FASTR3D (24), FRASS (25), SARA (26), ARTS (7) and RNA FRABASE (27) are available for RNA structural analysis. These servers employ different approaches, require different input types and produce different types of output. To our knowledge, the most similarly intended programs to COGNAC are the NASSAM and WebFR3D servers. The program WebFR3D (10) allows two modes of searching a database: a geometric search that finds all motifs similar to a user-specified query structure; and a symbolic search that finds all sets of nucleotides making user-specified interactions. The searchable structures for WebFR3D are currently limited to those already in the database because the program does not appear to allow for user-supplied structures to be analyzed.

The NASSAM server (9) takes a PDB formatted structure and searches it for a specific 3D arrangement of the RNA bases against its pattern database. Such an approach requires prior knowledge of the 3D arrangement for an interaction or motif and is therefore restricted when posed with generic queries that require, for example, all triples or quintuples in a structure to be listed. COGNAC, although able to search for triples that are defined by a 3D arrangement, does not find arrangements that are outside of the hydrogen-bonding parameters that have been specified. Although NASSAM has demonstrated a high degree of flexibility in extrapolating search patterns towards searching for novel orientations, it is still not able to accept a more general and less restrictive wildcard type query where, for example, the type and orientation of one or all the bases are not specifically defined. Searches that incorporate such wildcard options are advantageous for both the searching of novel base orientations and the searching of large hydrogen bonded base-base interaction clusters. Due to the obvious difficulties involved in theoretically predicting base interaction orientations for clusters in excess of three bases, the COGNAC querying capability presents an ideal solution to resolve this issue.

An accounting of the current database used revealed that sextuple interactions represented by the tree patterns sextuple types IV and V in Figure 2B have yet to be found in any existing reasonably high resolution X-ray crystallographic RNA structures in the PDB. This is not unsurprising for the type V sextuple pattern due to the number of interactions the central node is required to have and thus may probably not be sterically or chemically possible. Searching with COGNAC was able to retrieve the expected extended planar interactions consisting of mainly triples and quadruples (Figure 3A). Interactions that included bases orientated off plane to each other (Figure 3B) that were still able to form hydrogen bonds had been observed for the networks involving more than three bases. In the structure of a lysine riboswitch (PDB ID: 4erl, 4erj) (28), two adjacent quadruples were detected by COGNAC (Figure 3B, center panel). These adjacent quadruples were composed of different bases with one being an all guanine quadruple (Figure 3B, top panel) and the other comprising of two cytosines and two uracils (Figure 3B, bottom panel). Despite the base composition difference, both quadruples consisted of a similar arrangement where one base is connected to three other bases. This observation is also interesting due to the fact that the two separate quadruples are interconnected and therefore are components of a larger octuple interaction (Figure 3B, center panel).

SUMMARY

The COGNAC web server, when deployed to annotate existing RNA structures, is expected to further complement existing resources by bridging the gap between the complexity of RNA tertiary interactions to highly organized annotation schemes and 2D representations. From the testing and comparison of the currently available services, it is clear that the information and types of outputs provided by the different web servers do not overlap. To our knowledge, there are also no available servers that enable a user to compare the differences in the hydrogen bond interactions present between two very similar structures. This is an especially useful feature when attempting to detect the minute differences that can occur between RNA structures in different states such as between mutants and wild-types or between ligand bound and unbound states. In fact, the WebFR3D, NASSAM and COGNAC services can be seen as complementary to each other and can provide an array of tools to address different questions that can be posed for an RNA structure.

ACCESSION NUMBERS

PDB IDs: 6tna, 4erl and 4erj.

ACKNOWLEDGMENTS

We gratefully acknowledge the Genome Computing Centre of the Malaysia Genome Institute for providing the computational infrastructure for operating the COGNAC web server. We thank Mohd Noor Mat Isa, Hafiza Aida Ahmad and Sri Devan Appasamy for technical assistance with server operations. This paper is dedicated to the memory of Professor Peter J. Artymiuk who passed away on 19 February 2014.

FUNDING

Universiti Kebangsaan Malaysia [ICONIC-2013-007]; University Malaya High Impact Research MOE Grant [UM.C/625/1HIR/MoE/FCSIT/002 – Acc. No.: B222001]; Ministry of Education Malaysia [FRGS/1/2013/ST04/UKM/02/2]. Universiti Sains Malaysia postgraduate fellowship (to H.Y.H.); National Science Fellowship, Ministry of Higher Education, Malaysia (to N.N.). Funding for open access charge: University Malaya High Impact Research MOE Grant [UM.C/625/1HIR/MoE/FCSIT/002 – Acc. No.: B222001].

Conflict of interest statement. None declared.

REFERENCES

1. Klostermeier,D. and Millar,D.P. (2002) Energetics of hydrogen bond networks in RNA: hydrogen bonds surrounding G+1 and U42 are the major determinants for the tertiary structure stability of the hairpin ribozyme. *Biochemistry*, **41**, 14095–14102.
2. Plazanet,M., Fukushima,N. and Johnson,M. (2002) Modelling molecular vibrations in extended hydrogen-bonded networks, crystalline bases of RNA and DNA and the nucleosides. *Chem. Phys.*, **280**, 53–70.
3. Dingley,A.J., Masse,J.E., Feigon,J. and Grzesiek,S. (2000) Characterization of the hydrogen bond network in guanosine quartets by internucleotide 3hJ(NC)' and 2hJ(NN) scalar couplings. *J. Biomol. NMR*, **16**, 279–289.
4. Jucker,F.M., Heus,H.A., Yip,P.F., Moors,E.H. and Pardi,A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.*, **264**, 968–980.
5. Otero,R., Schock,M., Molina,L.M., Laegsgaard,E., Stensgaard,I., Hammer,B. and Besenbacher,F. (2005) Guanine quartet networks stabilized by cooperative hydrogen bonds. *Angew. Chem. Int. Ed. Engl.*, **44**, 2270–2275.
6. Chang,Y.F., Huang,Y.L. and Lu,C.L. (2008) SARSA: a web tool for structural alignment of RNA using a structural alphabet. *Nucleic Acids Res.*, **36**, W19–W24.
7. Dror,O., Nussinov,R. and Wolfson,H. (2005) ARTS: alignment of RNA tertiary structures. *Bioinformatic*, **21**(Suppl. 2), ii47–ii53.
8. Ferre,F., Ponty,Y., Lorenz,W.A. and Clote,P. (2007) DIAL: a web server for the pairwise alignment of two RNA three-dimensional structures using nucleotide, dihedral angle and base-pairing similarities. *Nucleic Acids Res.*, **35**, W659–W668.
9. Hamdani,H.Y., Appasamy,S.D., Willett,P., Artymiuk,P.J. and Firdaus-Raih,M. (2012) NASSAM: a server to search for and annotate tertiary interactions and motifs in three-dimensional structures of complex RNA molecules. *Nucleic Acids Res.*, **40**, W35–W41.
10. Petrov,A.I., Zirbel,C.L. and Leontis,N.B. (2011) WebFR3D—a server for finding, aligning and analyzing recurrent RNA 3D motifs. *Nucleic Acids Res.*, **39**, W50–W55.
11. Lescoute,A. and Westhof,E. (2006) Topology of three-way junctions in folded RNAs. *RNA* , **12**, 83–93.
12. Leontis,N.B. and Westhof,E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA*, **7**, 499–512.
13. Gutell,R.R., Power,A., Hertz,G.Z., Putz,E.J. and Stormo,G.D. (1992) Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods. *Nucleic Acids Res.*, **20**, 5785–5795.
14. Abu Almakarem,A.S., Petrov,A.I., Stombaugh,J., Zirbel,C.L. and Leontis,N.B. (2012) Comprehensive survey and geometric

classification of base triples in RNA structures. *Nucleic Acids Res.*, **40**, 1407–1423.

15. Sarver,M., Zirbel,C.L., Stombaugh,J., Mokdad,A. and Leontis,N.B. (2008) FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. *J. Math. Biol.*, **56**, 215–252.
16. Appasamy,S.D., Ramlan,E.I. and Firdaus-Raih,M. (2013) Comparative sequence and structure analysis reveals the conservation and diversity of nucleotide positions and their associated tertiary interactions in the riboswitches. *PLoS One*, **8**, e73984.
17. Firdaus-Raih,M., Harrison,A.M., Willett,P. and Artymiuk,P.J. (2011) Novel base triples in RNA structures revealed by graph theoretical searching methods. *BMC Bioinformatics*, **12**(Suppl. 13), S2–S15.
18. Gan,H.H., Pasquali,S. and Schlick,T. (2003) Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design. *Nucleic Acids Res.*, **31**, 2926–2943.
19. Bruno,I.J., Kemp,N.M., Artymiuk,P.J. and Willett,P. (1997) Representation and searching of carbohydrate structures using graph-theoretic techniques. *Carbohydr. Res.*, **304**, 61–67.
20. Deshpande,N., Addess,K.J., Bluhm,W.F., Merino-Ott,J.C., Townsend-Merino,W., Zhang,Q., Knezevich,C., Xie,L., Chen,L., Feng,Z. *et al.* (2005) The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Res.*, **33**, D233–D237.
21. McDonald,I.K. and Thornton,J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.*, **238**, 777–793.
22. Ullmann,J.R. (1976) An algorithm for subgraph isomorphism. *J. ACM*, **23**, 31–42.
23. Lorenz,R., Bernhart,S., Honer zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P. and Hofacker,I. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26–40.
24. Lai,C.E., Tsai,M.Y., Liu,Y.C., Wang,C.W., Chen,K.T. and Lu,C.L. (2009) FASTR3D: a fast and accurate search tool for similar RNA 3D structures. *Nucleic Acids Res.*, **37**, W287–W295.
25. Kirillova,S., Tosatto,S.C. and Carugo,O. (2010) FRASS: the web-server for RNA structural comparison. *BMC Bioinformatics*, **11**, 327–335.
26. Capriotti,E. and Marti-Renom,M.A. (2008) RNA structure alignment by a unit-vector approach. *Bioinformatics*, **24**, i112–i118.
27. Popenda,M., Szachniuk,M., Blazewicz,M., Wasik,S., Burke,E.K., Blazewicz,J. and Adamiak,R.W. (2010) RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinformatics*, **11**, 231–243.
28. Garst,A.D., Porter,E.B. and Batey,R.T. (2012) Insights into the regulatory landscape of the lysine riboswitch. *J. Mol. Biol.*, **423**, 17–33.